the globus® toolkit
www.globustoolkit.org

NMI

A Globus® Primer

What is the Grid and How Do I Use It?

Lee Liming
U. of Chicago and Argonne National Laboratory

Globus • NSF Middleware Initiative • TeraGrid

the globus® toolkit
www.globustoolkit.org

NMI

Approach

I. What is the Grid?
II. How are Grids built and used (today)?
III. What Grid software is available and what does it do?

BREAK (somewhere during part III)

IV. How have others succeeded?

COME CELEBRATE WITH US!

the globus® toolkit
www.globustoolkit.org

In appreciation of your support of all things Globus over the past decade, you are cordially invited to the Globus 10th Birthday Party.

When: Monday, September 11, 2006 - 7:00pm, immediately following Ian Foster's Globus State of the Union Keynote.

Where: The convention center concourse, in the center of the GlobusWORLD / GridWorld conference activity.

What: Food, drinks, music, friends and lots of fun!

the globus® toolkit
www.globustoolkit.org

NMI

I. What is the Grid?

History

- For years, a few whacky computer scientists have been trying to help other scientists use distributed computing.
  - Interactive simulation (climate modeling)
  - Very large-scale simulation and analysis (galaxy formation, gravity waves, battlefield simulation)
  - Engineering (parameter studies, linked component models)
  - Experimental data analysis (high-energy physics)
  - Image and sensor analysis (astronomy, climate study, ecology)
  - Online instrumentation (microscopes, x-ray devices, etc.)
  - Remote visualization (climate studies, biology)
  - Engineering (large-scale structural testing, chemical engineering)
- In these cases, the scientific problems are big enough that they require people in several organizations to collaborate and share computing resources, data, instruments.

the globus toolkit
www.globustoolkit.org
NMI

What Types of Problems?

- Your system administrators can't agree on a uniform authentication system, but you have to allow your users to authenticate once (using a single password) then use services on all systems, with per-user accounting.
- You need to be able to offload work during peak times to systems at other companies, but the volume of work they'll accept changes from day-to-day.

What Types of Problems?

- You and your colleagues have 6000 datasets from the past 50 years of studies that you want to start sharing, but no one is willing to submit the data to a centrally-managed storage system or database.
- You need to run 24 experiments that each use six large-scale physical experimental facilities operating together in real time.

Two Cardinal Rules of the Grid

- You can't rely on homogeneity.
  - Impossible to achieve in the real world.
  - STRATEGY - Plan on dealing with diverse systems and use mechanisms to *manage* heterogeneity.
- You can't rely on trust.
  - Severely limits participation.
  - STRATEGY - Provide a security model that can *express* complicated social networks.
  - STRATEGY - Use full disclosure when making requests (who is requesting, authorizing, and authenticating the request) and give service owners tools to *enforce* local policies.

Some Core Problems

- Heterogeneity
  - Too hard to keep track of authentication data (ID/password) across institutions
  - Too hard to monitor system and application status across institutions
  - Too many ways to submit jobs
  - Too many ways to store & access files and data
  - Too many ways to keep track of data
  - Too many opportunities to leave "dangling" resources lying around (robustness)
- Trust
  - Rigid use policies (authorization, QoS) and rigid application assumptions don't mix well.
  - Authorization needs to happen at many levels (communities, organizations, resource owners, etc.).
  - Complicated social structures exceed the abilities of simple authorization systems.

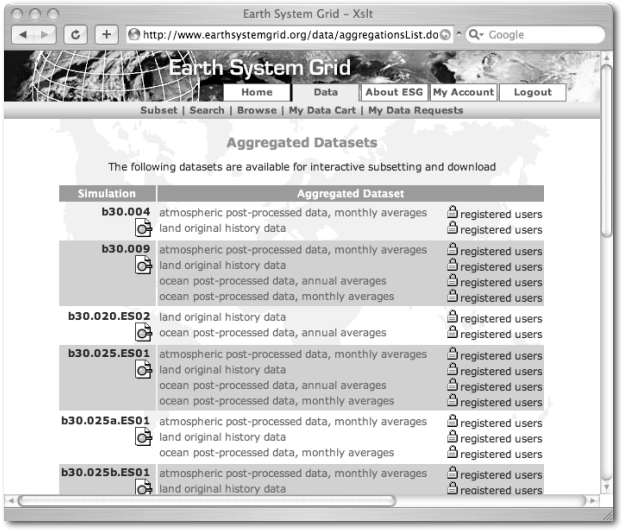
Challenging Applications

- The applications that Grid technology is aimed at are *not* easy applications!
  - The reason these things haven't been done before is because people believed it was too hard to bother trying.
  - If you're trying to do these things, you'd better be prepared for it to be challenging.
- Grid technologies are aimed at helping to overcome the challenges.
  - They solve some of the most common problems
  - They encourage standard solutions that make future interoperability easier
  - They were developed as parts of real projects
  - In many cases, they benefit from years of lessons from multiple applications
  - Ever-improving documentation, installation, configuration, training

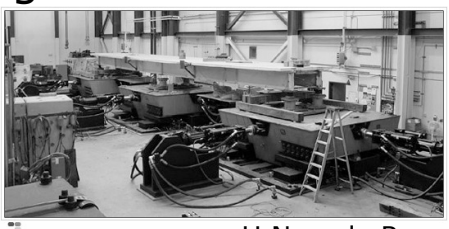
Earth System Grid

Goal:
Give climate scientists easier access to the distributed data and resources that they require to perform their research.

Developed new technologies for (1) creating and operating "filtering servers" capable of performing sophisticated analyses, and (2) delivering results to users.



Collaborative Engineering: NEES



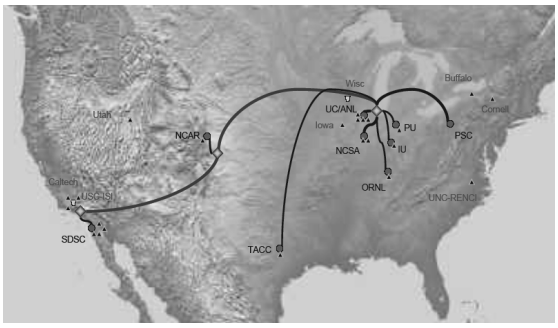
U.Nevada Reno

www.neesgrid.org

the globus toolkit
www.globustoolkit.org

NMI

NSF's TeraGrid*



A National Science Foundation Investment in Cyberinfrastructure

$100M 3-year construction (2001-2004)

$150M 5-year operation & enhancement (2005-2009)

- TeraGrid DEEP: Integrating NSF's most powerful computers (60+ TF)
  - 2+ PB Online Data Storage
  - National data visualization facilities
  - World's most powerful network (national footprint)
- TeraGrid WIDE Science Gateways: Engaging Scientific Communities
  - 90+ Community Data Collections
  - Growing set of community partnerships spanning the science community.
  - Leveraging NSF ITR, NIH, DOE and other science community projects.
  - Engaging peer Grid projects such as Open Science Grid in the U.S. as peer Grids in Europe and Asia-Pacific.
- Base TeraGrid Cyberinfrastructure: Persistent, Reliable, National
  - Coordinated distributed computing and information environment
  - Coherent User Outreach, Training, and Support
  - Common, open infrastructure services

* Slide courtesy of Ray Bair, Argonne National Laboratory

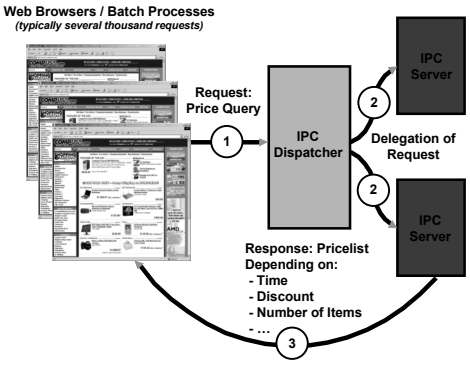
the globus toolkit
www.globustoolkit.org

NMI

Enterprise Workload Management: SAP AG R/3

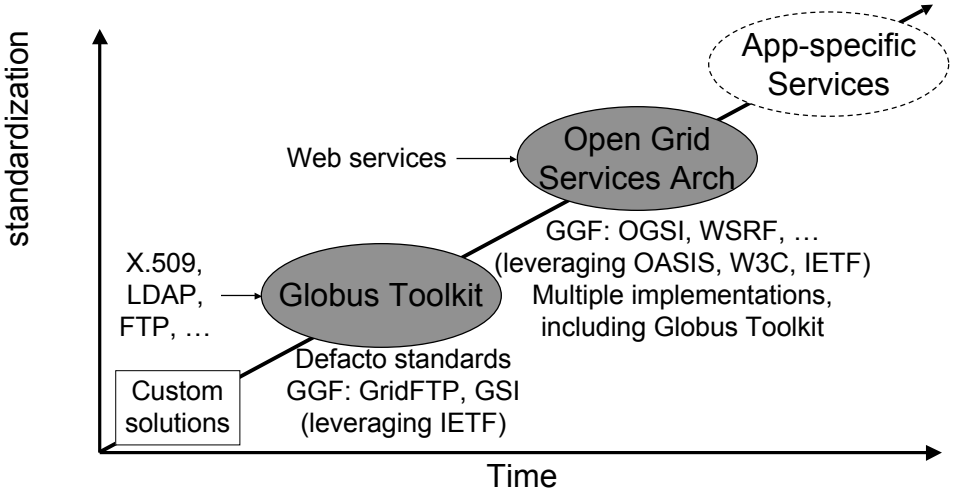
Goal: Lower cost of operation and improve performance of sales and pricing tools.

- Applied to R/3 products (CRM, SCM)
- Adapts to changing demand & resources
- Discovers and allocates resources as needed to meet specified response time goals



SAP AG R/3 Internet Pricing & Configurator (IPC)

Evolution of the Grid



Open Grid Services Architecture

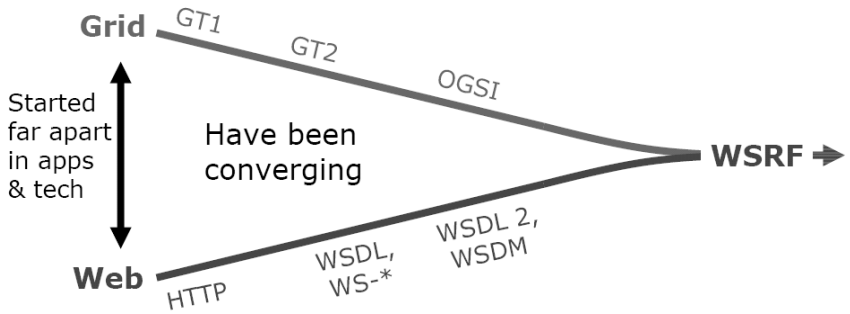
- Define a service-oriented architecture…
  - the key to effective virtualization
- …to address vital Grid requirements
  - AKA utility, on-demand, system management, collaborative computing, etc.
- …building on Web service standards.
  - extending those standards when needed

WSRF & WS-Notification

Patterns for Web services that enable Grid capabilities.

- Naming and bindings (basis for virtualization)
  - Every resource can be uniquely referenced, and has one or more associated services for interacting with it
- Lifecycle (basis for fault resilient state management)
  - Resources created by services following factory pattern
  - Resources destroyed immediately or scheduled
- Information model (basis for monitoring & discovery)
  - Resource properties associated with resources
  - Operations for querying and setting this info
  - Asynchronous notification of changes to properties
- Service Groups (basis for registries & collective svcs)
  - Group membership rules & membership management
- Base Fault type

Grid and Web Services Convergence



The definition of WSRF means that the Grid and Web services communities can move forward on a common base.

Why Not Just Use XML/SOAP?

- WSRF and WS-N *are* just XML and SOAP.
- WSRF and WS-N are just Web services.
- Benefits of following the specs:
  - These patterns represent best practices that have been learned in many Grid applications.
  - There is a community behind them.
  - Why reinvent the wheel?
  - Standards facilitate interoperability.

An "Ecosystem" of Grid Software

- There isn't a Grid software kit for everybody (yet).
  - Varying requirements
  - Experimentation and learning
  - Reluctance to invest in a "static" solution
- There are many tools that work well together.
  - Results of successful projects
  - Reusable solutions
- Implication: Integrate it yourself (for now).
  - Provides considerable flexibility
  - Requires expertise and effort
- Reminder: *These are ambitious applications!*

NMI

II. How Are Grids Built and Used Today?

NMI

Methodology

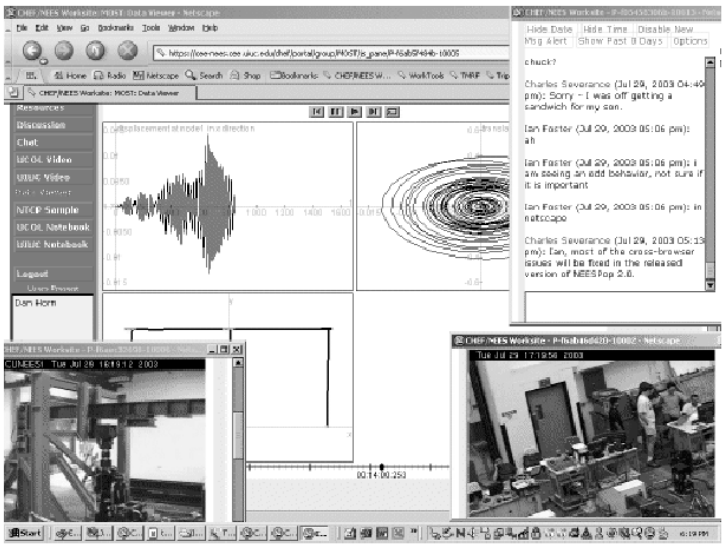
- Building a Grid system or application is currently an exercise in software *integration*.
  - Define user requirements
  - Derive system requirements or features
  - Survey existing components
  - Identify useful components
  - Develop components to fit into the gaps
  - Integrate the system
  - Deploy and test the system
  - Maintain the system during its operation
- This should be done iteratively, with many loops and eddies in the flow.

the globus toolkit
www.globustoolkit.org

NMI

What End Users Need

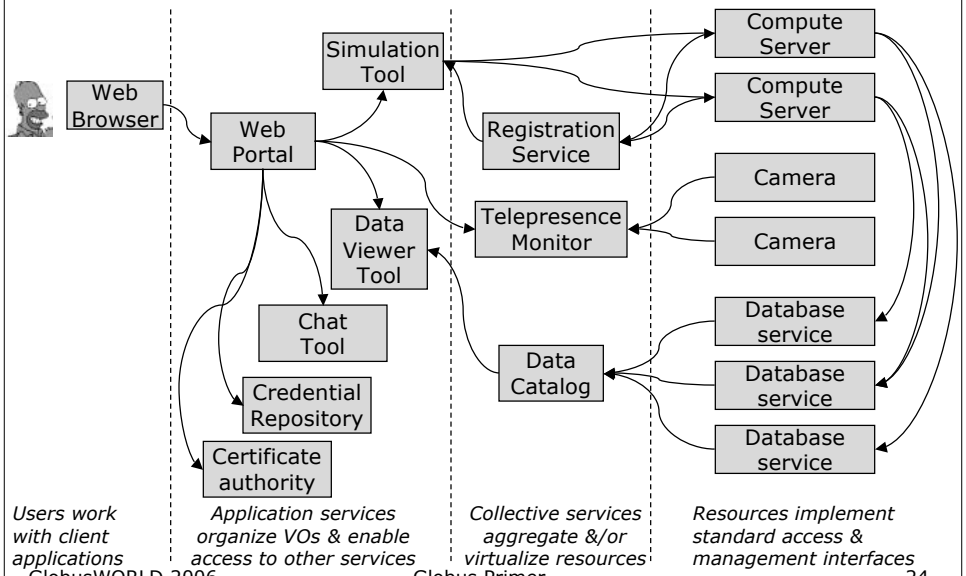
Secure, reliable, on-demand access to data, software, people, and other resources (ideally all via a Web Browser!)



the globus toolkit
www.globustoolkit.org

NMI

How it *Really* Happens



How it *Really* Happens

- Implementations are provided by a mix of
  - Application-specific code
  - "Off the shelf" tools and services
  - Tools and services from the Grid community (Globus + others using the same standards)
- Glued together by…
  - Application development
  - System integration

The Importance of Community

- All Grid technology is evolving rapidly.
  - Web services standards
  - Grid interfaces
  - Grid implementations
  - Grid resource providers (ASP, SSP, etc.)
- Community is important!
  - Best practices (OGF, OASIS, etc.)
  - Open source (Linux, Axis, Globus, etc.)
- Application of community standards is *vital*.
  - Increases leverage
  - Mitigates (a bit) effects of rapid evolution
  - Paves the way for future integration/partnership

What Is the Globus Toolkit?

- The Globus Toolkit is a collection of solutions to problems that frequently come up when trying to build collaborative distributed applications.
- Heterogeneity
  - To date (v1.0 - v4.0), the Toolkit has focused on *simplifying heterogenity* for application developers.
  - We are increasingly including more "vertical solutions" that implement typical application patterns.
- Security
  - The Grid Security Infrastructure (GSI) provides security mechanisms that operate at the service/community level, allowing collaborators to share resources without blind trust.
- Standards
  - Our goal has been to capitalize on and encourage use of existing standards (IETF, W3C, OASIS, OGF).
  - The Toolkit also includes reference implementations of new/proposed standards in these organizations.

Leveraging Existing and Proposed Standards

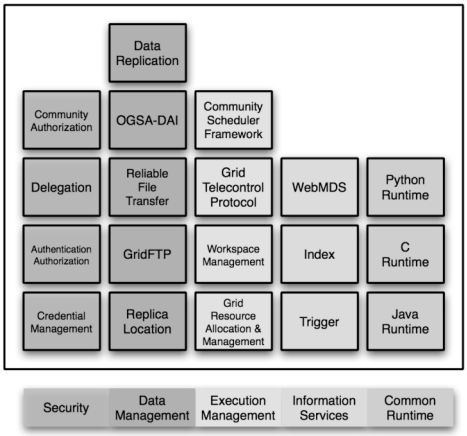
- SSL/TLS v1 (from OpenSSL) (IETF)
- X.509 Proxy Certificates (IETF)
- GridFTP v1.0 (OGF)
- WSRF and WS-N (OASIS)
- And others on the road to standardization:

  DAI, WS-Agreement, WSDL 2.0, WSDM, SAML, XACML

Areas of Competence

- "Connectivity Layer" Solutions
  - Service Management (WS Core)
  - Monitoring/Discovery (WS Core)
  - Security (GSI and WS-Security)
  - Communication (XIO)
- "Resource Layer" Solutions
  - Computing / Processing Power (GRAM)
  - Data Access/Movement (GridFTP, OGSA-DAI)
  - In development: Telecontrol (GTCP)
- "Collective Layer" Solutions
  - Data Management (RLS, DRS, RFT, OGSA-DAI)
  - Monitoring/Discovery (Index, Trigger, Archiver services)
  - Security (CAS, MyProxy)

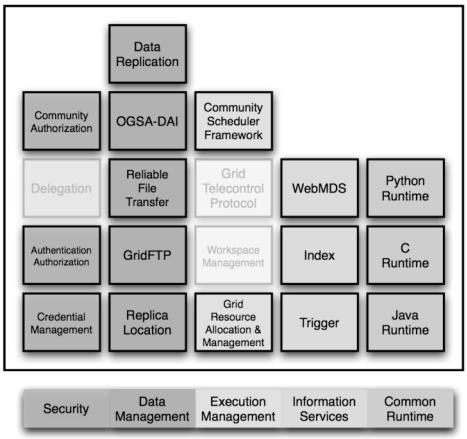
Globus Toolkit 4.0 Components



the globus toolkit
www.globustoolkit.org

NMI

Covered in This Tutorial



the globus toolkit
www.globustoolkit.org

NMI

What's In the Globus Toolkit?

- A Grid development environment
  - Develop new OGSA-compliant Web Services
  - Develop applications using Java or C/C++ Grid APIs
  - Secure applications using basic security mechanisms
- A set of basic Grid services
  - Job submission/management
  - File transfer (individual, queued)
  - Database access
  - Data management (replication, metadata)
  - Monitoring/Indexing system information
- Tools and Examples
- The prerequisites for many Grid community tools

How To Use the Globus Toolkit

- By itself, the Toolkit has surprisingly limited *end user* value.
  - There's very little user interface material there.
  - You can't just give it to end users (scientists, engineers, marketing specialists) and tell them to do something useful!
- The Globus Toolkit is useful to *application developers* and *system integrators*.
  - You'll need to have a specific *application or system* in mind.
  - You'll need to have the right *expertise*.
  - You'll need to set up prerequisite *hardware/software*.
  - You'll need to have a *plan*.

the globus toolkit
www.globustoolkit.org

III. What Grid Software Is Available And What Does It Do?

Security Tools

- Basic Grid Security Mechanisms
- Certificate Generation Tools
- Certificate Management Tools
  - Getting users "registered" to use a Grid
  - Getting Grid credentials to wherever they're needed in the system
- Authorization/Access Control Tools
  - Storing and providing access to system-wide authorization information

Basic Grid Security Mechanisms

- Basic Grid authentication and authorization mechanisms come in two flavors.
  - Pre-Web services
  - Web services
- Both are included in the Globus Toolkit, and both provide vital security features.
  - Grid-wide identities implemented as PKI certificates
  - Transport-level and message-level authentication
  - Ability to delegate credentials to agents
  - Ability to map between Grid & local identities
  - Local security administration & enforcement
  - Single sign-on support implemented as "proxies"
  - A "plug in" framework for authorization decisions

Basic Grid Security Mechanisms

- Basic security mechanisms are provided as libraries/classes and APIs.
  - Integrated with other GT tools and services
  - Integrated with many Grid community tools and services (and applications & systems)
- A few stand-alone tools are also included.

A Cautionary Note

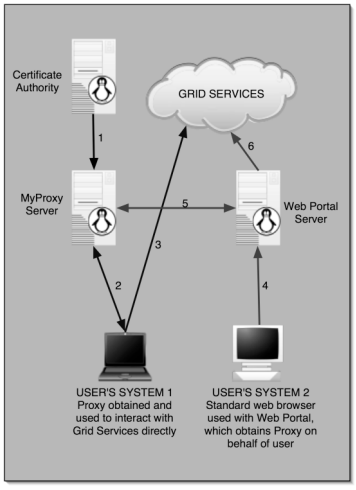
- Grid security mechanisms are tedious to set up.
  - If exposed to users, hand-holding is usually required.
  - These mechanisms can be *hidden entirely* from end users, but still used behind the scenes.
- These mechanisms exist for good reasons.
  - Many useful things can be done without Grid security.
  - It is unlikely that an ambitious project could go into production operation without security like this.
  - Most successful projects end up using Grid security, but using it in ways that end users don't see much.

Simple CA

- A convenient method of setting up a certificate authority (CA).
  - The Certificate Authority can then be used to issue certificates for users and services that work with GSI and WS-Security.
  - Simple CA is intended for operators of small Grid testing environments and users who are not part of a larger Grid.
- Most production Grids will not accept certificates that are not signed by a well-known CA, so the certificates generated by Simple CA will usually not be sufficient to gain access to production services.

MyProxy

- MyProxy is a remote service that stores user credentials.
  - Users can request proxies for local use on any system on the network.
  - Web Portals can request user proxies for use with back-end Grid services.
- Grid administrators can pre-load credentials in the server for users to retrieve when needed.
- MyProxy can be configured to use a built-in CA, using Kerberos or other PAM modules for authentication. (This eliminates the need to pre-store certificates.)
- Greatly simplifies certificate management!



KX.509 and KCA

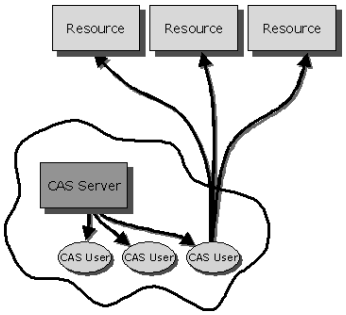
- Institutions that already have a Kerberos realm can use KX.509 and KCA to provide local users with Grid proxy certificates without using a Certificate Authority.
- When users authenticate with Kerberos, they may obtain proxy certificates in addition to their Kerberos tickets.
- KCA is a Kerberized certification service, and KX.509 is a Kerberized client that generates and stores proxy certificates.
- KX.509 and KCA create credentials for users, so remote sites must be configured to trust the local KCA service's certification authority.

PKINIT

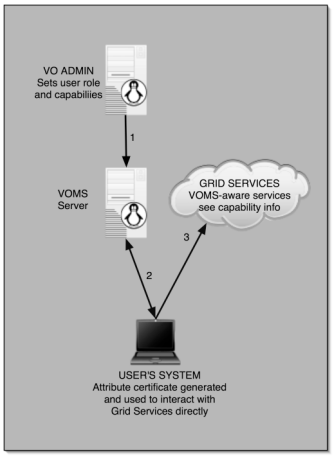
- PKINIT is a service that allows users to use Grid certificates to authenticate to a Kerberos realm.
- For sites that use Kerberized services (like AFS), this allows remote Grid users to obtain the necessary Kerberos tickets to use the site's local facilities properly.
- PKINIT replaces the Kerberos "klog" command and uses the user's Grid certificate to eliminate the need for a Kerberos passphrase.

CAS: Community Authorization Service

- CAS allows resource providers to specify course-grained access control policies in terms of communities as a whole.
- Fine-grained access control is delegated to the community.
- Resource providers maintain ultimate authority over their resources (including per-user control and auditing) but are spared most day-to-day policy administration tasks.



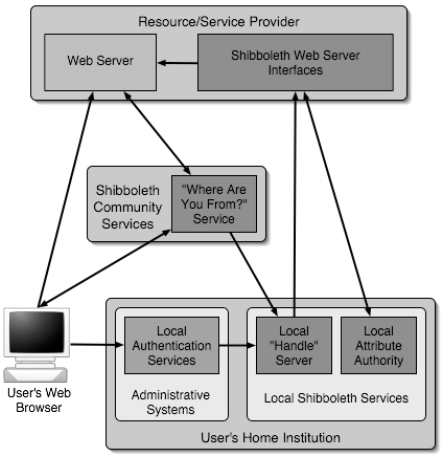
VOMS



- A community-level group membership system
- Database of user roles
  - Administrative tools
  - Client interface
- voms-proxy-init
  - Uses client interface to produce an attribute certificate (instead of proxy) that includes roles & capabilities signed by VOMS server
  - Works with non-VOMS services, but gives more info to VOMS-aware services
- Allows VOs to centrally manage user roles

Shibboleth

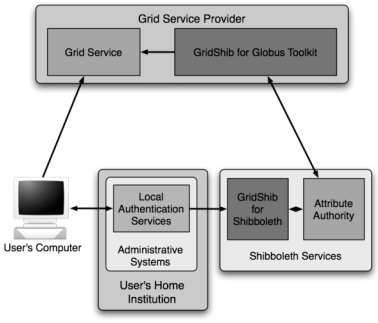
- Federated authorization and authentication management for Web-based services*
- Home institutions handle user authentication and role ("attribute") information
- Resource providers use role info for authorization
- Allows for user privacy (resource provider sees a "handle" from the home institution, which can protect identity)



* *There is ongoing work to extend Shibboleth capabilities to non-browser applications, including Grid systems*

GridShib

- Plugins for Globus and Shibboleth
- Allows Grid proxies to be used to look up Shibboleth attributes
  - User obtains Grid proxy from local authentication system (KX.509, MyProxy CA)
  - Authorization is based on user's proxy plus Shibboleth attributes



NMI

Break

Next up:
III. More of the Grid Ecosystem
IV. How Others Succeeded

NMI

Monitoring/Discovery Tools

- Basic OGSA Infrastructure Components
- Specialized Monitoring/Discovery Components
  - Specialized collection/monitoring agents
  - Viewing and display tools for showing system information for a variety of specialized purposes

OGSA Infrastructure Elements

- WS Core Monitoring Features
  - GT4 implements WS-ResourceProperties and WS-Notification.
- These features provide:
  - a standard interface for obtaining status and configuration information
  - a standard interface for clients to subscribe to particular information (i.e., notification)
  - the basis for registration and "grouping" of related services

Globus Index Service

- Provides a registry capability
  - Services register with the index service to make their presence known to other Grid components
  - Index service can also pro-actively subscribe to specific services based on configuration
    - Data, datatype, data provider information
- Provides a caching capability
  - Caches resource property values from registered services
- Indexes can be set up for a variety of uses, projects

Globus Trigger Service

- Triggers email alerts when pre-defined conditions are met
  - Especially useful for system status alerts in "production" operation situations
- Monitoring behavior
  - Subscribes to a set of resource properties
  - Runs a set of pre-configured tests on the resulting data streams to evaluate trigger conditions
  - When a trigger condition matches, sends email to the pre-configured address(es)

Archive Service

- Logs the values of resource property information over time
- Especially useful for auditing uptime/downtime for service agreement and/or diagnostic purposes
- Consumes resource property data
  - Subscribes to a set of resource properties
  - Stores resulting data in a database (Xindice)
- Provides access to stored data
  - Other consumers can contact database archive interface

WebMDS

- Web browser interface to resource properties and index services
  - Collects monitoring information from pluggable sources
    - By default, by using WSRF Resource Properties poll operations
    - MDS archiver plug in development to display historical information
  - Formats monitoring data (in XML) into HTML output using pluggable XSLT style sheets
- Runs in Tomcat

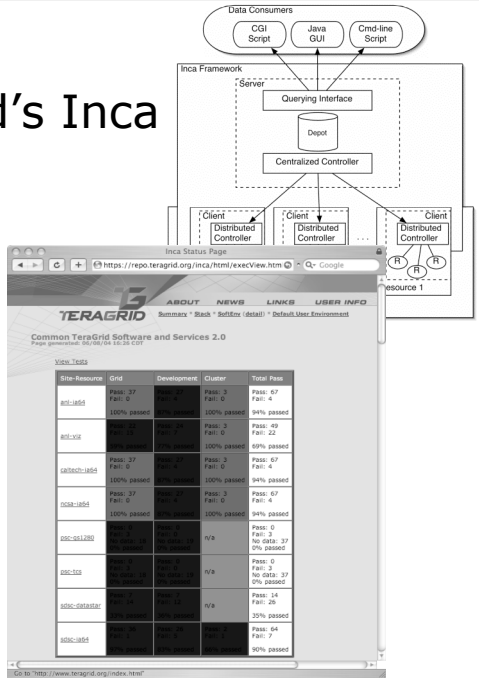
Ganglia Cluster Toolkit

- Ganglia is a toolkit for monitoring clusters and aggregations of clusters (hierarchically).
- Ganglia collects system status information and makes it available via a web interface.
- Ganglia status can be subscribed to and aggregated across multiple systems.
- Integrating Ganglia with MDS services results in status information provided in the proposed standard GLUE schema, popular in international Grid collaborations.

the globus toolkit
www.globustoolkit.org

TeraGrid's Inca

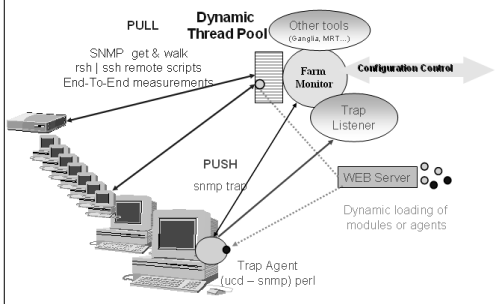
- Hierarchical Status Monitoring
  - Groups tests into logical sets
  - Supports many levels of detail and summarization
- Flexible, scalable architecture
  - Very simple reporter API
  - Can use existing test scripts (unit tests, status tools)
  - Hierarchical controllers
  - Several query/display tools



the globus toolkit
www.globustoolkit.org

NMI

MonALISA



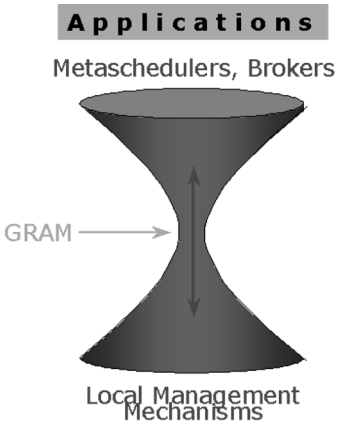
- Supports system-wide, distributed monitoring with triggers for alerts
  - Java/JINI and Web services
  - Integration with Ganglia, queuing systems, Condor, etc.
  - Client interfaces include Web and WAP
  - Flexible registration and proxy mechanisms support look-ups and firewalls

Computing/Processing Tools

- Workflow Managers
  - Organize and coordinate task execution within a complicated application
  - Often coordinates data movement and task execution
- Metaschedulers
  - Optimize use of distributed compute pools
- Virtual Data Tools
  - Manage the trade-off between data storage and processing power

GRAM - Basic Job Submission and Control Service

- A uniform service interface for remote job submission and control
  - Includes file staging and I/O management
  - Includes reliability features
  - Supports basic Grid security mechanisms
  - Available in Pre-WS and WS
- GRAM is *not* a scheduler.
  - No scheduling
  - No metascheduling/brokering
  - Often used as a front-end to schedulers, and often used to simplify metaschedulers/brokers



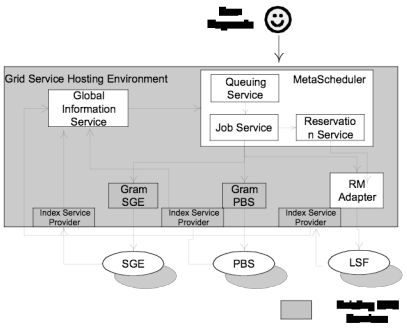
Grid-enabled Schedulers

- Scheduling systems that are easily integrated with GRAM via plug-ins
- Wide variety of capabilities (supported models)
- Wide variety of support (commercial vs. open source)
- Note that GRAM can be used as *either* an interface to a scheduler *or* the interface that a scheduler uses to submit a job to a resource.

- Condor
- OpenPBS
- Torque
- PBSPro
- Sun Grid Engine
- Platform LSF

Platform CSF

- An open source implemenation of OGSA-based metascheduler for VOs.
  - Supports emerging WS-Agreement spec
  - Supports GT GRAM
  - Uses GT Index Service
- Fills in gaps in existing resource management picture
  - Integrated with Platform LSF and Platform Multicluster
  - Included in GT4

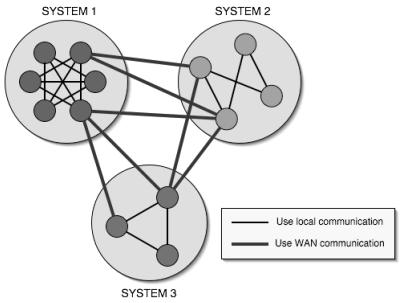


Condor-G, DAGman

- Condor-G and DAGman address many workflow challenges for Grid applications.
  - Managing sets of subtasks
  - Getting the tasks done reliably and efficiently
  - Submitting to Grid resources via GRAM
  - Checkpointing and Migration

MPICH-G2

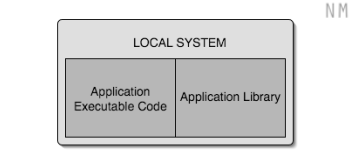
- A Grid-enabled implementation of the MPI v1.1 framework
- For applications that have concurrent jobs that communicate with each other
- Chooses communication method (shared memory, local high-performance interconnect, IP, etc.) based on location of source & destination
- Uses Grid for security, job submission, WAN communication, etc.



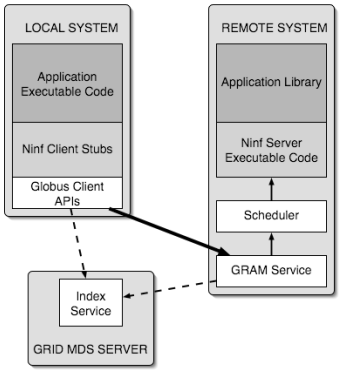
MPICH-G2 automatically chooses the best communication method for messages based on the location of each subjob and other factors.

Ninf-G

- Grid-enabled RPC implementation
- Allows existing libraries to be used as RPC calls rather than local calls
- Implements GGF's GridRPC specification



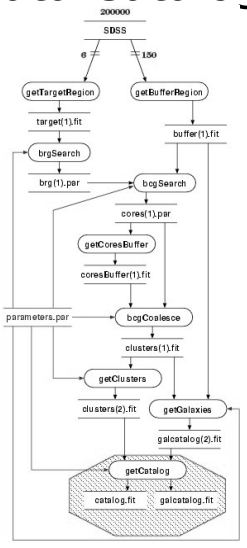
Without Ninf: Library linked directly to application



With Ninf: Library linked to Ninf server code and application linked to Ninf client stubs

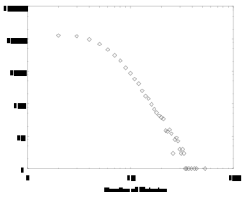
Virtual Data Catalog

- Captures both logical and physical steps in a data analysis process.
  - Transformations (logical)
  - Derivations (physical)
- Builds a catalog.
- Results can be used to "replay" analysis.
  - Generation of DAG (via Pegasus)
  - Execution on Grid
- Catalog allows introspection of analysis process.



Sloan Survey Data

Galaxy cluster size distribution

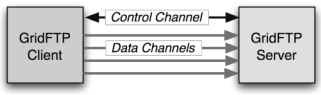


Data Tools

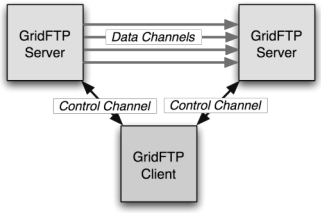
- Virtual Data Tools
  - Manage the trade-off between data storage and processing power (already covered)
- Movement/Transfer Tools
  - Interfaces that meet specialized application or user needs
  - "Last mile" integration to specialized storage systems
- Optimization Tools
  - Help optimize the use of storage systems for specialized user communities

GridFTP

- A high-performance, secure data transfer service optimized for high-bandwidth wide-area networks
  - FTP with extensions
  - Uses basic Grid security (control and data channels)
  - Multiple data channels for parallel transfers
  - Partial file transfers
  - Third-party (direct server-to-server) transfers
- OGF recommendation GFD.20

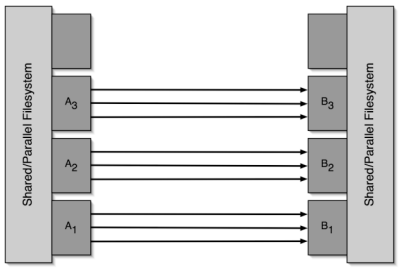


Basic Transfer
One control channel, several parallel data channels



Third-party Transfer
Control channels to each server, several parallel data channels between servers

Striped GridFTP



- GridFTP supports a striped (multi-node) configuration.
  - Establish control channel with one node
  - Coordinate data channels on multiple nodes
  - Allows use of many NICs in a single transfer
- Requires shared/parallel filesystem on all nodes.
  - On high-performance WANs, aggregate performance is limited by filesystem data rates.

globus-url-copy

- Command-line client for GridFTP servers
  - Text interface
  - No "interactive shell" (single command per invocation)
- Many features
  - Grid security, including data channel(s)
  - HTTP, FTP, GridFTP
  - Server-to-server transfers
  - Subdirectory transfers and lists of transfers
  - Multiple parallel data channels
  - TCP tuning parameters
  - Retry parameters
  - Transfer status output

UberFTP

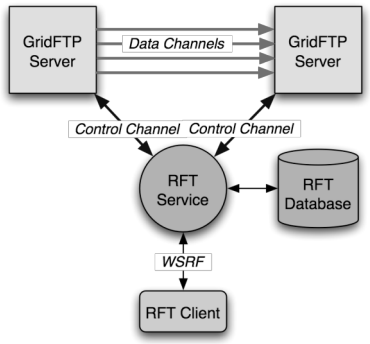
- UberFTP is an interactive (text prompt) client for GridFTP.
- Supports more features than NCFTP
  - Parallelism
  - Server-to-server transfers

GSI-SCP/SFTP

- GSI-OpenSSH is a version of OpenSSH that supports Grid authentication.
  - Remote terminal (gsi-ssh)
  - Remote Copy (gsi-scp)
  - Secure FTP (gsi-sftp)
- More familiar to many users than GridFTP.
- Doesn't take advantage of GridFTP's capabilities (parallelism, partial files, third-party transfers, etc.)

RFT - File Transfer Queuing

- A WSRF service for queuing file transfer requests
  - Server-to-server transfers
  - Checkpointing for restarts
  - Database back-end for failovers
- Allows clients to request transfers and then "disappear"
  - No need to manage the transfer
  - Status monitoring available if desired



OGSA-DAI

- OGSA interface for accessing XML and relational data stores
- Implements the GGF DAIS WG standard (in progress)

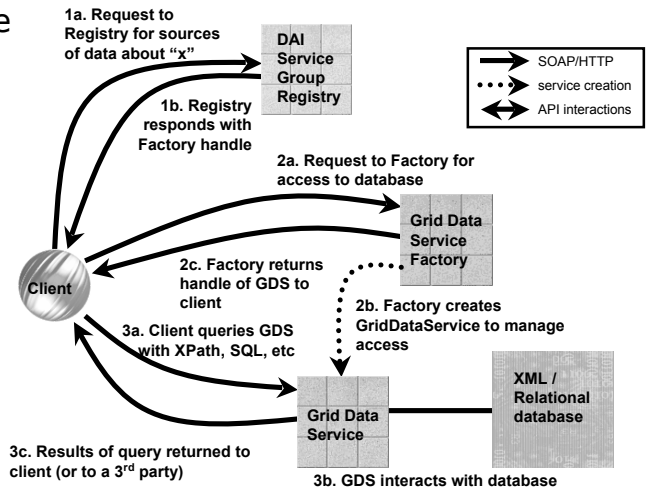


Figure courtesy of Malcolm Atkinson and Rob Baxter, UK eScience Center

MCS - Metadata Catalog Service

- A stand-alone metadata catalog service
  - OGSA service interface
  - Stores system-defined and user-defined attributes for logical files/objects
  - Supports manipulation and query
- Integrated with OGSA-DAI
  - OGSA-DAI provides metadata storage
  - When run with OGSA-DAI, basic Grid authentication mechanisms are available

RLS - Replica Location Service

- A distributed system for tracking replicated data
  - Consistent local state maintained in Local Replica Catalogs (LRCs)
  - Collective state with relaxed consistency maintained in Replica Location Indices (RLIs)
- Performance features
  - Soft state maintenance of RLI state
  - Compression of state updates
  - Membership and partitioning information maintenance

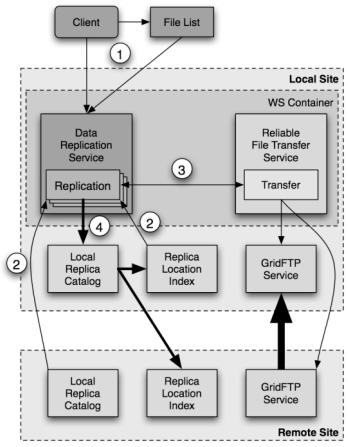
Note:

- RLS (developed by Globus Alliance and the DataGrid Project) replaces earlier components in the Globus Toolkit 2.x.

the globus toolkit
www.globustoolkit.org

NMI

Data Replication Service (DRS)



- Pull-mode File Replicator
  1. Clients add files to replication list
  2. RLS used to locate original copies
  3. RFT/GridFTP used to make local copies
  4. RLS updated with new locations
- Integrated with GT4
  - Uses WSRF-compliant services written in Java.
  - Works with RLS, RFT, and GridFTP services in GT4.
- Tech preview in GT4

the globus toolkit
www.globustoolkit.org

NMI

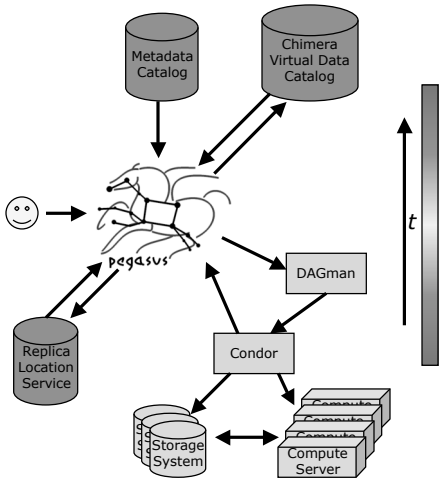
DataCutter

- Provides segmenting and processing capabilities for large (shared) datasets
  - Filters - Services that can be "placed" (executed) on computational resources near storage system to transform the data in some way between endpoints
  - Streams - Filters input and output data via streams, which can be unidirectional or bidirectional (pipe)
- Integrated with GSI (to secure streams), GRAM (to place filters), and MDS (to monitor and discover resources for filter placement)

Virtual Data System (VDS)

Workflow implemented using virtual data and Grid processing capabilities

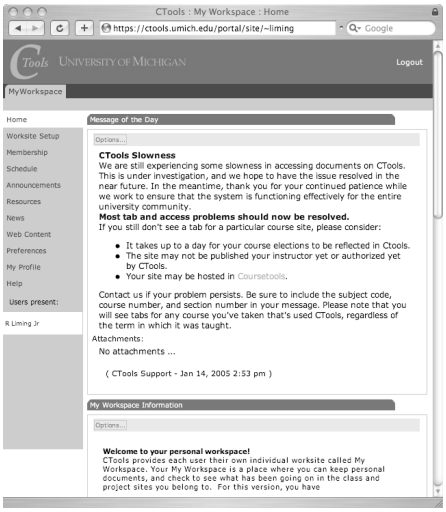
- Uses Metadata to convert user request to logical data sources
- Obtains abstract workflow info from a VDC (catalog)
- Uses replication data to locate physical files
- Delivers concrete workflow instructions to DAGman
- Executes using Condor
- Publishes new replication and derivation data in RLS and VDC



Collaboration Tools

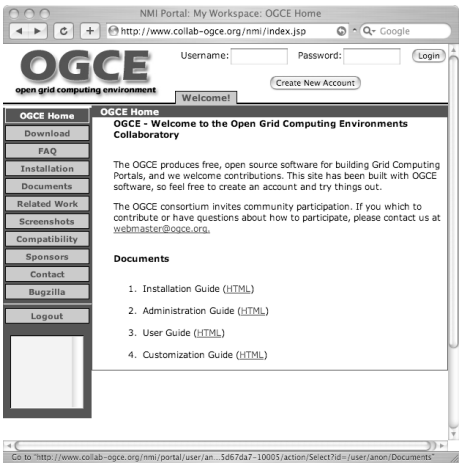
- Web Portals
  - Tools for building web interfaces that provide access to system/application capabilities
- High-function Virtual Meeting Spaces
  - Tools for facilitating collaborative work in realtime with lots of gadgets

Sakai



- Flexible web-based system for supporting distributed learning and collaborative work
- Allows users and groups to establish shared web spaces and add content via a browser interface
- Emphasizes JSR-168 portlet specification and localization
- Highly extensible, with support for JetSpeed, Velocity, and other portal interfaces

Open Grid Computing Environment (OGCE)



- Provides standard compliant portlet components for building Web portals. E.g.,
  - MyProxy
  - GridPort
  - GT services (GRAM, GridFTP, MDS, etc.)
  - Java CoG
- Provides a "quick start" for building Grid-enabled portals.

Access Grid

- Provides seamless group-to-group collaboration spaces for groups that are not co-located
  - Immersive audio (sounds like "everyone is here")
  - Multiple multicast video streams (multipoint "everyone sees everyone")
  - Can integrate other video sources into the space
  - Display walls common, but not required
- "Virtual spaces" (like channels) allow people to find each other
- Ideally, groups can work together without thinking about the technology.
- Increasingly Grid-enabled

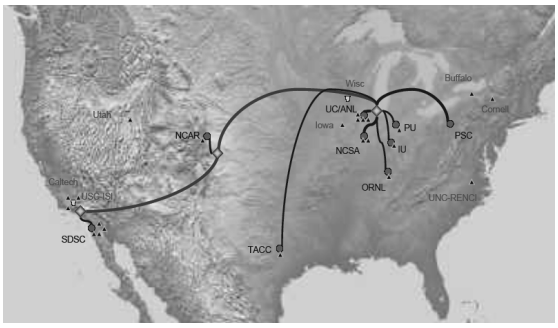
the globus toolkit
www.globustoolkit.org

IV. How Others Have Succeeded

the globus toolkit
www.globustoolkit.org

NMI

NSF's TeraGrid*



A National Science Foundation Investment in Cyberinfrastructure

$100M 3-year construction (2001-2004)

$150M 5-year operation & enhancement (2005-2009)

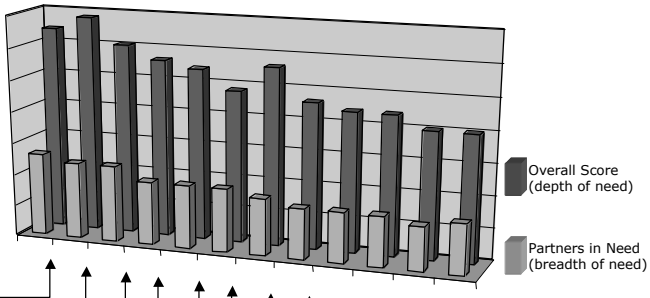
- TeraGrid DEEP: Integrating NSF's most powerful computers (60+ TF)
  - 2+ PB Online Data Storage
  - National data visualization facilities
  - World's most powerful network (national footprint)
- TeraGrid WIDE Science Gateways: Engaging Scientific Communities
  - 90+ Community Data Collections
  - Growing set of community partnerships spanning the science community.
  - Leveraging NSF ITR, NIH, DOE and other science community projects.
  - Engaging peer Grid projects such as Open Science Grid in the U.S. as peer Grids in Europe and Asia-Pacific.
- Base TeraGrid Cyberinfrastructure: Persistent, Reliable, National
  - Coordinated distributed computing and information environment
  - Coherent User Outreach, Training, and Support
  - Common, open infrastructure services

* Slide courtesy of Ray Bair, Argonne National Laboratory

the globus toolkit
www.globustoolkit.org

NMI

TeraGrid User Priorities*

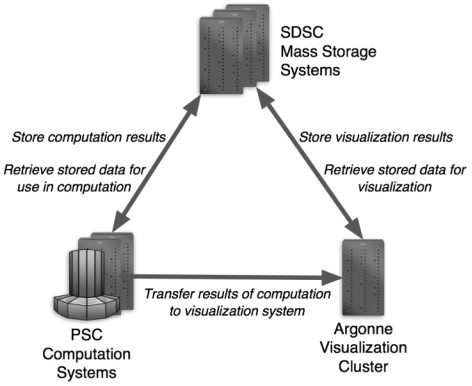


Remote File Read/Write
High-Performance File Transfer
Coupled Applications, Co-scheduling
Grid Portal Toolkits
Grid Workflow Tools
Batch Metascheduling
Global File System
Client-Side Computing Tools
Batch Scheduled Parameter Sweep Tools
Advanced Reservations

* Slide courtesy of Charlie Catlett, TeraGrid Project Director

Data Transfer on TeraGrid

- TeraGrid encourages sites to offer unique resources that address different needs.
- Because of this, users may need to move data from one TeraGrid system to another.
- Data transfer must be made as efficient as possible, especially from a "person time" perspective.



The Challenge

- Early TeraGrid users were disappointed with data transfer rates between sites.
  - 270 Mbps over a 30 Gbps link
  - Requires knowledge of server and network configuration
  - High-performance tools (e.g., globus-url-copy) not as friendly as low-performance tools (scp)
- For some users, managing transfers was also an issue that consumed too much "person time."
- CHALLENGE: Provide an easy-to-use tool that provides high-performance data transfer when using default settings, reducing human supervision time as much as possible.

Issues - Technical

- Networking hardware
  - Ordinary software tools (scp) use single TCP streams, which can't consume the bandwidth of a NIC (1 Gbps).
  - Most TeraGrid NICs are either 100 Mbps or 1 Gbps, with very few 2 Gbps. These can't consume bandwidth of a 30 Gbps link.
  - REQUIREMENT: The tool must employ both parallelism (within a host) and striping (multiple hosts).
  - REQUIREMENT: Hosts must be provisioned.
- Filesystems
  - Local filesystem performance (disk I/O, local connectivity) is also a limiting factor.
  - REQUIREMENT: Parallel filesystems are needed.
- Complexity
  - Specialized storage systems at sites may provide unique interfaces. This increases application/user complexity.
  - REQUIREMENT: Provide a uniform interface to storage systems ("virtualization") wherever possible.

Issues - Social (1)

- Who has information?
  - User/application knows about local path and remote path. System administrators know about servers and network configuration/properties.
  - REQUIREMENT: The tool should get details about servers and network parameters to be used in transfers from administrators, not users.
- How do expectations get set?
  - Users know about published specs, not historical experience (30 Gbps, not 300 Mbps). This raises expectations far beyond reality.
  - REQUIREMENT: The solution should include a way to capture historical experience and publish it to users.

Issues - Social (2)

- How many transfers are typical?
  - Big files may require restarts to complete. Large numbers of small files must be "shepherded."
  - REQUIREMENT: The tool should provide an automated management capability to reduce user "shepherding."
- Security (Authentication/Authorization)
  - Users typically have different accounts at different sites. (This is true for "application" accounts as well.) It's hard to keep track of this information.
  - REQUIREMENT: The tool should not require users or applications to remember their various local accounts at each site.

Requirements Summary

- The tool must employ both parallelism (within a host) and striping (multiple hosts).
- Hosts must be provisioned.
- Parallel filesystems are needed.
- Provide a uniform interface to storage systems ("virtualization") wherever possible.
- The tool should get details about servers and network parameters to be used in transfers from administrators, not users.
- The solution should include a way to capture historical experience and publish it to users.
- The tool should provide an automated management capability to reduce user "shepherding."
- The tool should not require users or applications to remember their various local accounts at each site.

Proposed Solution

- Allocate a set of nodes at each site to serve collectively as a striped GridFTP server.
  - Access to shared filesystems
  - Ideally, fast NICs
- Provide a command that offers a simple, scp-like interface.
  - Use administrator-provided configuration data to redirect to striped servers when possible and add optimized network parameters based on endpoints.
  - Use globus-url-copy and rft clients to perform the transfers.

TGCP - TeraGrid Copy

- Applies a set of transformation rules to source and destination.
  - Local admin supplies the rules.
  - Adds host/port and appropriate path information, puts into GridFTP URL format.
- When source/dest sites are identified, add network tuning parameters based on a table maintained by administrators.
- Invoke either g-u-c or rft to perform the transfer. Pass through any command line options.

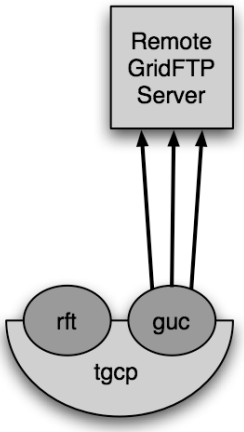
TGCP User Interface

- SCP-style source and destination
  - host:path
- Two options
  - -big - Use striped transfer
  - -rft - Manage the transfer
- If source is a directory, RFT will transfer the full contents of the directory.

```
% tgcp sim-slices.xml \
? tg-login.sdsc.edu:/u/liming/viz
```

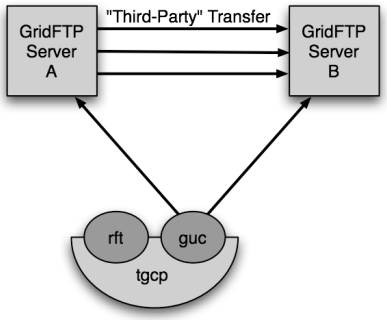

Scenario #1 - Non-shared Local Filesystem

- If the local GridFTP servers can't get to the local file, tgcp uses g-u-c as a GridFTP client.
  - Parallelism and network tuning parameters help to optimize transfer.



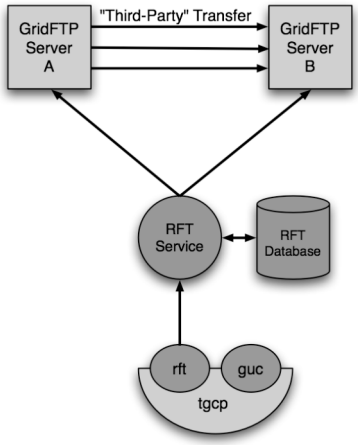
Scenario #2 - High-Performance Transfer

- If the file is accessible to servers at both ends, tgcp can invoke a third party transfer.
  - Striping can be performed as well as parallelism.
  - Servers typically have high-performance NICs.



Scenario #3 - Managed High-perf. Transfer

- If there are many files to be transferred (e.g., a directory) or the user wants to "fire and forget", tgcp uses rft.
  - Requests stored in a persistent database.
  - Failure recovery
  - RFT can use channel caching for even better performance with small files.

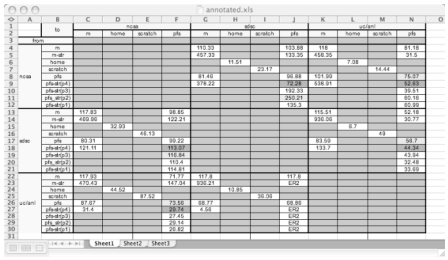


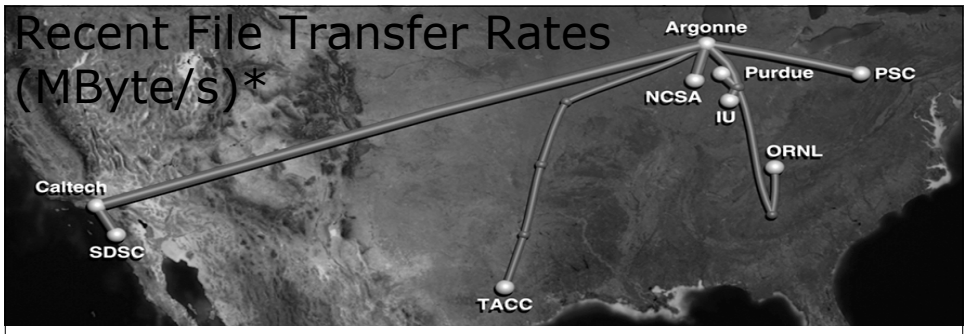
the globus toolkit
www.globustoolkit.org

NMI

Deployment and Monitoring

- TGCP team produced a TGCP "package" for sites to deploy locally.
- Documentation prepared for users.
- GT4 GridFTP, g-u-c, and tgcp deployed at all eight TeraGrid sites in "experimental" mode.
  - Non-default ports
  - New hosts
  - Softenv keys for user paths
- Speed runs are conducted roughly weekly with large files, and results shared.





	NCSA	SDSC	ANL	PSC	Purdue	Indiana	ORNL
NCSA		121	55	56	22	50	56
SDSC	198		57	32	26	69	191
ANL	44	23		9	20	41	35
PSC	69	84	91		22	38	53
Purdue	35	55	31	187		37	51
Indiana	91	102	70	15	26		64
ORNL	123	116	60	61	24	66	

* Slide courtesy of Ray Bair, data produced by Jaebum Kim, NCSA, Sept. 2005

Results - TeraGrid

- TGCP is available, deployed, and monitored.
  - Not quite x10 performance improvement
    - ~270 Mbps -> ~1.8 Gbps
  - Much simpler to use (don't need to specify parameters or shepherd as much)
- GT4 GridFTP and RFT are deployed.
  - Compatibility with other grid systems
  - Better support
- Forged relationship with NMI/GRIDS.
  - Teams have worked together on something real.
- Remaining work is in provisioning hosts, improving local filesystem performance.
  - TGCP will be able to use whatever the system offers.

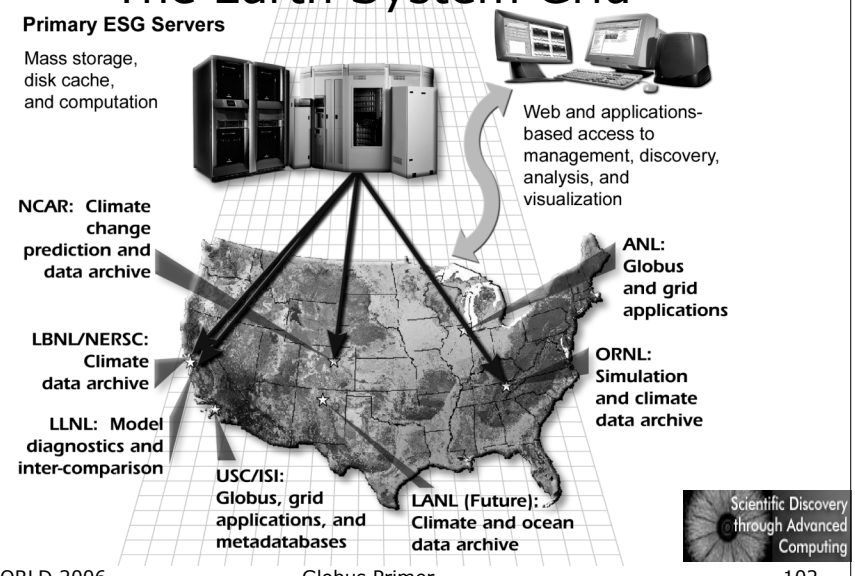
Results - Science

- Results so far are inconclusive.
  - TGCP is not yet declared "production."
  - Applications are still adjusting, haven't yet gained experience sufficient to offer judgments.
- TGCP software is available.
  - NMI/GRIDS offers it to other projects.

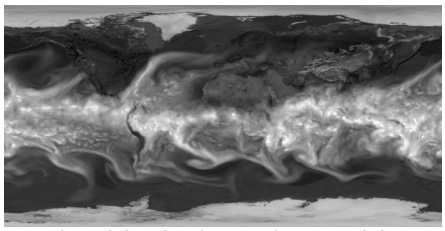
A Few TGCP Experiences

- TeraGrid resource providers (sites) are funded separately from integration team.
  - We needed to convince them to deploy the tools.
  - We relied on each site to properly configure the tools (GridFTP and TGCP config files).
  - The monitoring service and reports were key to obtaining a full deployment.
- Software is just the beginning.
  - TGCP removed the "tooling" bottleneck; new bottlenecks have appeared.
  - Performance is limited (now) by hardware provisioning at each site: servers, NICs, and local filesystems. *Dynamic provisioning may be key to solving this problem.*
  - Local filesystem performance is critical.

The Earth System Grid



ESG Project Goals



- Improve productivity/capability for the simulation and data management team (data producers).
- Improve productivity/capability for the research community in analyzing and visualizing results (data consumers).
- Enable broad multidisciplinary communities to access simulation results (end users).
- The community needs an integrated "cyberinfrastructure" to enable smooth *workflow* for *knowledge development*: compute platforms, collaboration & collaboratories, data management, access, distribution, and analysis.

The Challenge

- ESG is a distributed system that genuinely requires Grid-style distributed authentication.
- ESG is used by scientists who don't need to be bothered with certificates.
- CHALLENGE: Provide Grid security for the system but do it in such a way that end users don't have to manage certificates themselves.

Issues - Social

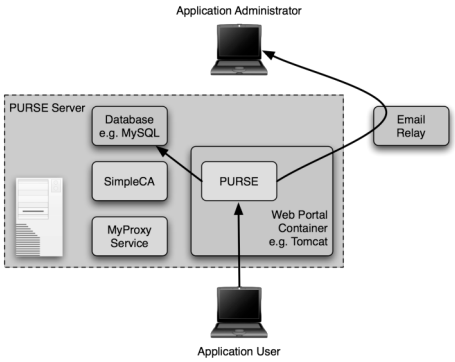
- Ease of Use
  - ESG users shouldn't have to manage their own certificates.
  - It's too complicated, intrusive.
  - They don't do it well (securely).
- Support
  - Certificate management generates a lot of user support work.
- Use cases
  - Most ESG users are data "readers", not "writers."
  - Data producers and project funders want to know who the users are (registration), but access control among registered users is not a major requirement.

Issues - Technical

- Distributed System
  - ESG has four major data centers, each with its own security system.
  - Users should not have to keep track of four sets of credentials and know when to use each.
  - The ESG web portal needs users' credentials to perform work on their behalf, so a secure mechanism for doing that is important.
- Integration
  - ESG uses GridFTP and GRAM to meet other system requirements, so GSI has to be supported.

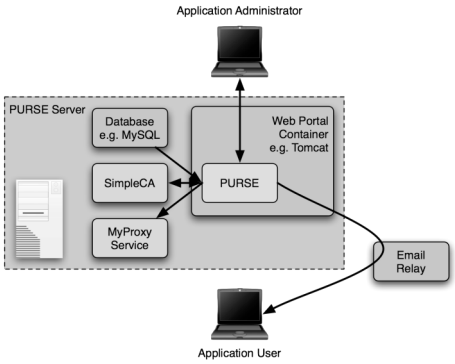
Scenario 1 - User Registration

- The user fills out the registration web page, establishes an ID/password, and the information is stored in a database.
- The administrator is sent email.



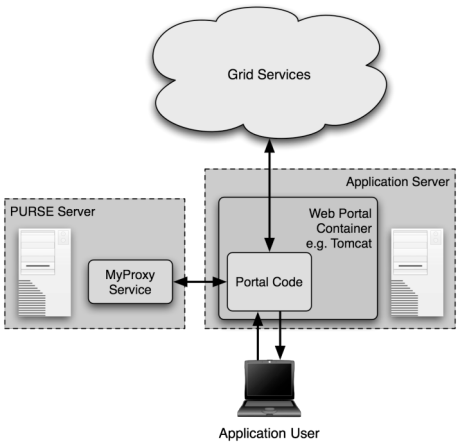
Scenario 2 - Administrator Approval

- Administrator visits the registration website and retrieves the registration data.
- If the administrator approves the request, PURSE uses SimpleCA to generate a certificate and stores it in MyProxy.
- The user is sent email.

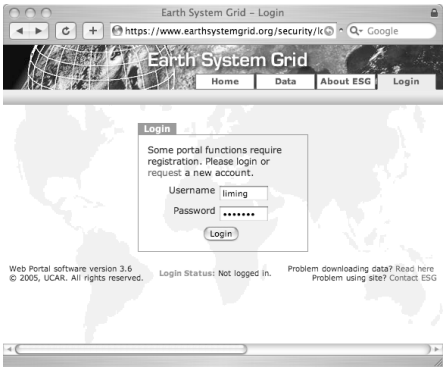


Scenario 3 - User Login

- The user logs into the application website using the ID/password established during registration.
- The application obtains a proxy using MyProxy.
- The application uses the proxy to authenticate to Grid services.



Results - ESG



- Four data centers (LBNL, LLNL, NCAR, ORNL)
- 700 registered users as of May 2005
- Two major datasets are available, with associated metadata
- Work underway to add IPCC datasets as they are produced

Results - Science

- ESG allows 700+ people to work with climate model datasets.
- PURSE is available from NMI/GRIDS Center.
  - Generic version for re-use
  - Allows users to import existing credentials
  - Supported by NMI program
- GAMA is available from SDSC.
  - Portlet implementation hosted by GridSphere
  - Allows sharing by multiple portal applications
  - Currently used by GEON and BIRN projects

A Few PURSE Lessons

- It is possible (and desirable) to hide Grid security from users.
  - Online repositories are one way to do this.
  - Others options include online CAs (e.g., KCA and KX.509).
- Requirements and use cases are important.
  - Need to know exactly what the community concerns are: what needs to be protected.
  - Need to clearly identify roles.
- Generalizing to PURSE was not trivial.
  - New requirements (e.g., credential import)
  - Documentation and usability testing
  - We still missed important requirements: JSR-168 compliant portlets and support for non-Grid accounts.

NMI

Conclusions

the globus toolkit
www.globustoolkit.org

NMI

Lessons Learned

- The Globus Toolkit has useful stuff in it.
- To do anything significant, a lot more is needed.
  - The Grid community (collectively) has many useful tools that can be reused!
  - System integration expertise is mandatory.
- OGSA and community standards (GGF, OASIS, W3C, IETF) are extremely important in getting all of this to work together.
- There's much more to be done!

Continue Learning

- Visit the Globus Alliance website at: www.globus.org
- Read the books:
  - *The Grid: Blueprint for a New Computing Infrastructure (2nd edition)*
  - *Globus Toolkit 4: Programming Java Services*
  - *Grid Computing: The Savvy Manager's Guide*
- Talk to others who are using the Toolkit: gt-user@globus.org (subscribe first)
- Participate in standards organizations: OGF, OASIS, W3C, IETF
- Attend GlobusWORLD, an annual event

Q: Do you see a fun & exciting career in my future?

Magic 8 Ball:
Signs point to **YES.**

Say YES to Great Career Opportunities!

Argonne NATIONAL LABORATORY

SOFTWARE ENGINEER/ARCHITECT
Mathematics and Computer Science Division, Argonne National Laboratory
The Grid is one of today's hottest technologies, and our team in the Distributed Systems Laboratory (www.mcs.anl.gov/dsl) is at the heart of it. Send us a resume through the Argonne site (www.anl.gov/Careers/), requisition number **MCS-310886.**

THE UNIVERSITY OF CHICAGO

SOFTWARE DEVELOPER
Computation Institute, University of Chicago
Join a world-class team developing pioneering eScience technologies and applications. Apply using the University's online employment application (http://jobs.uchicago.edu/, click "Job Opportunities" and search for requisition numbers **072817 and 072442**).

See our posting on the GlobusWorld job board - talk to the Globus team.

the globus® toolkit
www.globustoolkit.org

Meet the Developers Session

Globus Alliance Booth (152A-P7)

Tuesday, September 12

8:00 - 9:00am	"Java WS Core and Security (C, Java)" -- Olle Mulmo, Jarek Gawor, Rachana Ananthakrishnan
11:30 -12:30pm	"RLS" -- Rob Schuler, Ann Chervenak
12:30 -1:30pm	"MDS" -- Mike D'arcy, Laura Pearlman
3:00 - 4:00pm	"Resource Management (GRAM, Virtual Workspaces and Dynamic Accounts)" -- Stu Martin, Peter Lane, Tim Freeman, Kate Keahey
6:00 - 7:00pm	"C WS Core" -- Joe Bester
7:00 - 8:00pm	"Python WS Core" -- Joshua Boverhof

Wednesday, September 13

8:00 - 9:00am	"GridShib" -- Von Welch, Ton Scavo, Tim Freeman
11:30 - 12:30pm	"GT Installation and Administration" -- Charles Bacon
12:30 - 1:30pm	"MyProxy" -- Jim Basney
3:00 - 4:00pm	"GridFTP, XIO, RFT" -- John Bresnahan, Ravi Madduri

COME CELEBRATE WITH US!

the globus® toolkit
www.globustoolkit.org

In appreciation of your support of all things Globus over the past decade, you are cordially invited to the Globus 10th Birthday Party.

When: Monday, September 11, 2006 - 7:00pm, immediately following Ian Foster's Globus State of the Union Keynote.

Where: The convention center concourse, in the center of the GlobusWORLD / GridWorld conference activity.

What: Food, drinks, music, friends and lots of fun!